Springer Series in Operations Research and Financial Engineering

For further volumes:
http://www.springer.com/series/3182

John R. Birge • François Louveaux

Introduction to Stochastic Programming

Second Edition

Springer

John R. Birge
Booth School of Business
University of Chicago
5807 South Woodlawn Avenue
Chicago, Illinois 60637
USA
john.birge@chicagobooth.edu

François Louveaux
Department of Business Administration
University of Namur
Rempart de la Vierge 8
B-5000, Namur
Belgium
francois.louveaux@fundp.ac.be

ISSN 1431-8598
ISBN 978-1-4614-0236-7 e-ISBN 978-1-4614-0237-4
DOI 10.1007/978-1-4614-0237-4
Springer New York Dordrecht Heidelberg London

Library of Congress Control Number: 2011929942

Mathematics Subject Classification (2010): 37N40, 46N10, 49L20, 49Mxx (all), 49N30, 49N15, 90-01, 90B50, 90C05, 90C06, 90C15, 90C39

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

To Richard and Joelle,
Sebastien, Jérôme, Quentin, and Géraldine.

Preface

Since the publication of the first edition of this book, we have been encouraged by the growing interest in stochastic programming and its application in a variety of areas, including routine use in many industries from transportation and logistics to finance and energy. We have also been heartened by the many new methodological and theoretical advances within the field. In this second edition, we have attempted to capture aspects of both recent applications and models as well as new practically relevant methods and theory. As in the first edition, our primary goal is to provide students and other readers with an appreciation of how to build uncertainty into an optimization model, what differences in decisions might result from recognizing the presence of uncertainty, and how and what kinds of models are amenable to solution. We have focused the second edition on satisfying these main objectives while also uncovering basic research questions to give beginning researchers a foundation upon which to build more in-depth knowledge.

To help make the relevant issues in modeling, solving, and analyzing stochastic programs more evident, we have incorporated more examples than in the first edition so that the each of the main modeling, solution, and analysis processes are illustrated with a detailed example. We have also added many exercises whose solutions provide additional insights into stochastic programming concepts and tools. Many of these exercises assume the availability of software to solve basic linear and nonlinear optimization models and to construct algorithmic procedures involving matrix operations. Since we view completing these exercises as a key part of understanding the material, instructors should ensure that students have adequate programming skills to implement the methods described in the book.

Besides additional examples and exercises throughout the book, we have reorganized the material to improve the logical flow and to eliminate unnecessary or complicating issues before explaining the most practically relevant material. Specific changes in the second edition include the following:

- a new section (Section 1.5) and routing example in Chapter 1;
- a worked-out modeling exercise (Section 2.8) and a section on risk modeling and robust formulation (Section 2.9 in Chapter 2;

- re-arrangement and simplification of the material in Chapter 3 to emphasize basic model characteristics and illustrate them with examples;
- complete re-organization and combination of Chapters 5 and 6 into a new Chapter 5 that unifies the treatment of cutting-plane methods and again provides additional examples;
- an additional section on Lagrangian multistage methods in Chapter 6 (formerly Chapter 7);
- a completely re-organized version of Chapter 7 (formerly Chapter 8) including new methods and review material on combinatorial optimization;
- additional examples in Chapter 8 (formerly Chapter 9) including bounds on loss probabilities in loan portfolios;
- re-organization of Chapter 9 (formerly Chapter 10) to place practical methods earlier and to include a new section on Monte Carlo methods for probabilistic constraints;
- re-organization of Chapter 10 (formerly Chapter 11) to include new sections on scenario generation, multistage sampling methods, and approximate dynamic programming methods;
- removal of the short chapter (formerly Chapter 12) on a capacity expansion case study.

We anticipate that classes would follow much of the same sequence as we suggested for the first edition, but, with the increased availability of software to implement methods, we recommend that instructors include more computational exercises and additional modeling projects to fit students' interests. Any course should again start with the first two chapters to provide the application and modeling context. Depending on student interest, a typical class would generally include Chapters 3, 4, and Sections 5.1, 5.2, and 5.5 to present the most typical types of methods. For basic approximations, a modeling-focused class could focus on the main techniques in Chapters 8, 9, and 10 (for dynamic models), while a theoretically-oriented class might emphasize the analytical results in those chapters. A more computationally focussed class might emphasize the remainder of Chapter 5 plus Chapters 6 and 7.

We wish to thank the many people who sent us comments and suggestions about the first edition of the book and the numerous students we have worked with and all those who have helped us see stochastic programming from a fresh perspective every time we encounter something new. Among the many who have contributed, we thank Michael Dempster, Michel Gendreau, Maarten van der Vlerk, and Bill Ziemba. Thanks are also due to Martine Van Caeneghem for her patient typing of the modifications in Namur. We also again thank Fonds National de la Recherche Scientifique, the National Science Foundation, as well as the U.S. Department of Energy, and the University of Chicago Booth School of Business for their financial support.

In our first edition, we finished the preface with special thanks to our wives, Pierrette and Marie, to whom our book was dedicated. These thanks are more than ever very much present in our hearts. Now, we also want to express our proudness and joy of having such great children. We have thus decided to dedicate this second edition to them. We may thus expect that the third edition will be dedicated to our grandchildren, although the timing of this edition and the number of lines needed for this future dedication remain unknown.

Chicago, Illinois, USA John R. Birge
Namur, Belgium François Louveaux

Preface to the First Edition

According to a French saying "Gérer, c'est prévoir," which we may translate as "(The art of) Managing is (in) foreseeing." Now, probability and statistics have long since taught us that the future cannot be perfectly forecast but instead should be considered random or uncertain. The aim of stochastic programming is precisely to find an optimal decision in problems involving uncertain data. In this terminology, *stochastic* is opposed to *deterministic* and means that some data are random, whereas programming refers to the fact that various parts of the problem can be modeled as linear or nonlinear mathematical programs. The field, also known as *optimization under uncertainty*, is developing rapidly with contributions from many disciplines such as operations research, economics, mathematics, probability, and statistics. The objective of this book is to provide a wide overview of stochastic programming, without requiring more than a basic background in these various disciplines.

Introduction to Stochastic Programming is intended as a first course for beginning graduate students or advanced undergraduate students in such fields as operations research, industrial engineering, business administration (in particular, finance or management science), and mathematics. Students should have some basic knowledge of linear programming, elementary analysis, and probability as given, for example, in an introductory book on operations research or management science or in a combination of an introduction to linear programming (optimization) and an introduction to probability theory.

Instructors may need to add some material on convex analysis depending on the choice of sections covered. We chose not to include such introductory material because students' backgrounds may vary widely and other texts include these concepts in detail. We did, however, include an introduction to random variables while modeling stochastic programs in Section 2.1 and short reviews of linear programming, duality, and nonlinear programming at the end of Chapter 2. This material is given as an indication of the prerequisites in the book to help instructors provide any missing background. In the Subject Index, the first reference to a concept is where it is defined or, for concepts specific to a single section, where a source is provided.

In our view, the objective of a first course based on this book is to help students build an intuition on how to model uncertainty into mathematical programs, which changes uncertainty brings into the decision process, what difficulties uncertainty may bring, and what problems are solvable. To begin this development, the first section in Chapter 1 provides a worked example of modeling a stochastic program. It introduces the basic concepts, without using any new or specific techniques. This first example can be complemented by any one of the other proposed cases of Chapter 1, in finance, in multistage capacity expansion, and in manufacturing. Based again on examples, Chapter 2 describes how a stochastic model is formally built. It also stresses the fact that several different models can be built, depending on the type of uncertainty and the time when decisions must be taken. This chapter links the various concepts to alternative fields of planning under uncertainty.

Any course should begin with the study of those two chapters. The sequel would then depend on the students' interests and backgrounds. A typical course would consist of elements of Chapter 3, Sections 4.1 to 4.5, Sections 5.1 to 5.3 and 5.7, and one or two more advanced sections of the instructor's choice. The final case study may serve as a conclusion. A class emphasizing modeling might focus on basic approximations in Chapter 9 and sampling in Chapter 10. A computational class would stress methods from Chapters 6 to 8. A more theoretical class might concentrate more deeply on Chapter 3 and the results from Chapters 9 to 11.

The book can also be used as an introduction for graduate students interested in stochastic programming as a research area. They will find a broad coverage of mathematical properties, models, and solution algorithms. Broad coverage cannot mean an in-depth study of all existing research. The reader will thus be referred to the original papers for details. Advanced sections may require multivariate calculus, probability measure theory, or an introduction to nonlinear or integer programming. Here again, the stress is clearly in building knowledge and intuition in the field. Mathematical results are given so long as they are either basic properties or helpful in developing efficient solution procedures. The importance of the various sections clearly reflects our own interests, which focus on results that may lead to practical applications of stochastic programming.

To conclude, we may use the following little story. An elderly person, celebrating her one hundredth birthday, was asked how she succeeded in reaching that age. She answered, "It's very simple. You just have to wait."

In comparison, stochastic programming may well look like a field of young impatient people who not only do not want to wait and see but who consider waiting to be suboptimal. We realize how much patience was needed from our friends and colleagues who encouraged us to write this book, which took us much longer than expected. To all of them, we are extremely thankful for their support. The authors also wish to thank the Fonds National de la Recherche Scientifique and the National Science Foundation for their financial support. Both authors are deeply grateful to the people who introduced us to the field, George Dantzig, Roger Wets, Jacques

Drèze, and Guy de Ghellinck. Our special thanks go to our wives, Pierrette and Marie, to whom we dedicate this book.

Ann Arbor, Michigan John R. Birge
Namur, Belgium François Louveaux

Contents

Notation

The following describes the major symbols and notations used in the text. To the greatest extent possible, we have attempted to keep unique meanings for each item. In those cases where an item has additional uses, they should be clear from context. We include here only notation used in more than one section. Additional notation may be needed within specific sections and is explained when used.

In general, vectors are assumed to be columns with transposes to indicate row vectors. This yields $c^T x$ to denote the inner product of two n-vectors, c and x. We reserve prime ($'$) for first derivatives with respect to time (e.g., $f' = df/dt$).

Vectors in primal programs are represented by lowercase Latin letters while matrices are uppercase. Dual variables and certain scalars are generally Greek letters. Superscripts indicate a stage while subscripts indicate components followed by realization index. Boldface indicates a random quantity. Expectations of random variables are indicated by a bar ($\bar{\boldsymbol{\xi}}$), μ, or ($\mathrm{E}(\boldsymbol{\xi})$). We also use the bar notation to denote sample means in Chapter 9.

Equations are numbered consecutively in the text by section and number within the section (e.g., (1.2) for Section 1, Equation 2). For references to chapters other than the current one, we use three indices: chapter, section, and equation, (e.g., (3.1.2) for Chapter 3, Section 1, Equation 2). Exercises are given at the end of sections (or subsections in the cases of Sections 3.2 and 5.1) and are referenced in the same manner as equations. All other items (figures, tables, declarations, examples) are labeled consecutively through the entire chapter with a single reference (e.g., Figure 1) if within the current chapter and chapter and number if in a different chapter (e.g., Figure 3.1 for Chapter 3, Figure 1).

Symbol	Definition
$^+$	Superscript indicates the positive part of a real (i.e., $a^+ = \max(a,0)$) or unrestricted variable (e.g., $y = y^+ - y^-, y^+ \geq 0, y^- \geq 0$) and its objective coefficients (e.g., q^+), subscript as non-negative values in a set (e.g., $\Re_+$) or the right-limit ($F^+(t) = \lim_{s \downarrow t} F(s)$)
$^-$	Superscript indicates the negative part of a real (i.e., $a^- = \max(-a,0)$) or unrestricted variable (e.g., $y = y^+ - y^-, y^+ \geq 0, y^- \geq 0$) and its objective coefficients (e.g., q^-) or the left-limit ($F^-(t) = \lim_{s \uparrow t} F(s)$)
*	Indicates an optimal value or solution (e.g., x^*)
$^{0\ \hat{}\ \prime\ \sim}$	Indicate given nonoptimal values or solutions (e.g., x^0 , $\hat{x}$, x' , $\tilde{x}$)
0	Zero matrix (subscripts denote dimension when present)
$\mathbf{1}_X$	Indicator function of set X
a	Ancestor scenario, real value or vector
A	First-stage matrix (e.g., $Ax = b$), also used to indicate an event or subset, $A \in \mathscr{A} \subset \Omega$
$\mathscr{A}$	Collection of subsets
b	First-stage right-hand side (e.g., $Ax = b$)
B	Matrix, basis submatrix, Borel sets, or index set of a basis
$\mathscr{B}$	Collection of subsets (notably Borel sets)
c	First-stage objective ($c^T x$), t -th stage objective ($(c^t(\omega))^T x^t$) or real vectors
C	Matrix or index set of continuous variables
d	Right-hand side of a feasibility cut in the L-shaped method, a demand, or real vector
D	Left-hand side vector of a feasibility cut in the L-shaped method, a matrix, a set, or an index set of discrete variables
$\mathscr{D}$	Set of descendant scenarios
e	Exponential, right-hand side of an optimality cut in the L-shaped method, an extreme point, or the unit vector ($e^T = (1, \ldots, 1)$)
E	Mathematical expectation operator, left-hand side vector of an optimality cut in the L-shaped method, or an event
f	Function (usually in an objective ($f(x)$ or $f_i(x)$) or a density
F	Cumulative probability distribution

Symbol	Definition
g	Function (usually in constraints ($g(x)$ or $g_j(x)$))
h	Right-hand side in second-stage ($Wy = h - Tx$), also $h^t(\omega)$ in multistage problems
H	Number of stages (horizon) in multistage problems
i	Subscript index of functions (f_i) or vector elements (x_i , x_{ij})
I	Identity matrix or index set ($i \in I$)
j	Subscript index of functions (g_j) or vector elements (y_j , y_{ij})
J	Matrix or index set
k	Index of a realization of a random vector ($k = 1, \ldots, K$)
K	Feasibility sets (K_1, K_2) or total number of realizations of a discrete random vector
$\mathscr{K}$	Number of realizations or sample paths in a scenario tree with $\mathscr{K}^t$ nodes at stage t
l	Index, lower bound on a variable, or Lagrangian function
L	The L-shaped method, objective value lower bound, or real value
m	Number of constraints (m_1, m_2) or number of elements ($i = 1, \ldots, m$)
n	Number of variables (n_1, n_2) or number of elements ($i = 1, \ldots, n$)
N	Set, normal cone, normal distribution, or number of random elements
p	Probability of a random element (e.g., p_k $= \mathrm{P}(\boldsymbol{\xi} = \xi_k)$) or matrix of probabilities
P	Probability of events (e.g., $\mathrm{P}(\boldsymbol{\xi} \leq 0)$)
q	Second-stage objective vector ($q^T y$)
Q	Second-stage (multistage) value function with random argument ($Q(x, \boldsymbol{\xi})$ or $Q^t(x^t, \xi^t)$)
$\mathscr{Q}$	Second-stage (multistage) expected value value (recourse) function ($\mathscr{Q}(x)$ or $\mathscr{Q}^t(x^t)$)
r	Revenue or return in examples, real vector, or index
$\Re$	Real numbers
R	Matrix or set
s	Scenario or index

Symbol	Definition
S	Set or matrix
t	Superscript stage or period index for multistage programs ($t = 1,\ldots,H$), a real-valued parameter, or an index
T	Technology matrix ($Wy = h - Tx$ or $T^{t-1}(\omega)(x)$); as a superscript, the transpose of a matrix or vector
u	General vector, upper-bound vector, or expected shortage
U	Objective value upper bound
v	Variable vector or expected surplus
V	Set, matrix or an operator
w	Second-stage decision vector in some examples
W	Recourse matrix ($Wy = h - Tx$)
x	First-stage decision vector or multistage decision vector (x^t)
X	First-stage feasible set ($x \in X$) or t th stage feasible set (X^t)
y	Second-stage decision vector
Y	Second-stage feasible set ($y \in Y$)
z	Objective value ($\min z = c^T x + \cdots$)
Z	Integers
α	Real value, vector, or probability level with probabilistic constraints
β	Real value or vector
γ	Real value or function
δ	Real value or function
ε	Real value
ζ	Random variable
η	Real value or random variable
θ	Lower bound on $\mathscr{Q}(x)$ in the L-shaped method
κ	Index
λ	Dual multiplier, parameter in a convex combination, or measure
μ	Expectation (used mostly in examples of densities) or a parameter for non-negative multiples
ν	Algorithm iteration index (sometimes also the number of samples in Monte Carlo sampling algorithms)
$\boldsymbol{\xi}$	Random vector (often indexed by time, $\boldsymbol{\xi}^t$) with realizations as ξ (without boldface)
Ξ	Support of the random vector $\boldsymbol{\xi}$
π	Dual multiplier

Symbol	Definition
Π	Product, projection operator, or aggregated problem dual multiplier
ρ	Dual multiplier or discount factor
σ	Dual multiplier, standard deviation, or σ -field
Σ	Summation or covariance matrix
τ	Possible right-hand side in bundles or index of time
ϕ	Function in computing the value of the stochastic solution or a measure
Φ	Function, cumulative distribution of standard normal
$\emptyset$	Empty set
χ	Tender or offer from first to second period ($\chi = Tx$)
ψ	Second stage value function defined on tenders and with random argument, $\psi(\chi, \xi(\omega))$)
Ψ	Expected second stage value function defined on tenders, $\Psi(\chi)$)
ω	Random event ($\omega \in \Omega$)
Ω	Set of all random events